OSP-8939-42
US

日 本 国 特 許 庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日
Date of Application: １９９９年　２月２５日

出 願 番 号
Application Number: 平成１１年特許願第０４９２１５号

出 願 人
Applicant (s): 日本電信電話株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

２０００年　１月２８日

特許庁長官
Commissioner,
Patent Office

近藤隆彦

【書類名】　特許願

【整理番号】　NTTH107182

【提出日】　平成11年 2月25日

【あて先】　特許庁長官　殿

【国際特許分類】　H04L 12/00

【発明の名称】　トラヒック観測装置、データグラム転送システムおよびデータグラム転送方法

【請求項の数】　20

【発明者】

【住所又は居所】　東京都新宿区西新宿三丁目１９番２号　日本電信電話株式会社内

【氏名】　清水　敬司

【発明者】

【住所又は居所】　東京都新宿区西新宿三丁目１９番２号　日本電信電話株式会社内

【氏名】　栗本　崇

【特許出願人】

【識別番号】　000004226

【氏名又は名称】　日本電信電話株式会社

【代理人】

【識別番号】　100064908

【弁理士】

【氏名又は名称】　志賀　正武

【手数料の表示】

【予納台帳番号】　008707

【納付金額】　21,000円

【提出物件の目録】

【物件名】　明細書　1

【物件名】　図面　1

【物件名】　　要約書　1

【包括委任状番号】　9701417

【プルーフの要否】　　要

【書類名】　明細書

【発明の名称】　トラヒック観測装置、データグラム転送システムおよびデータグラム転送方法

【特許請求の範囲】

【請求項１】　現在よりあらかじめ決められた期間前までの、ユーザが網内に送信または網から受信したデータグラムに関するトラヒック情報を観測するトラヒック観測機能と、

該トラヒック観測機能により得られたトラヒック情報からデータグラムの送出に対する網の評価により得られるプレファレンス値を計算するプレファレンス値計算機能と、

該プレファレンス値計算機能により計算されたプレファレンス値を評価値として前記送信されるデータグラムのヘッダに書き込むプレファレンス値挿入機能とを備えたことを特徴とするトラヒック観測装置。

【請求項２】　ユーザ端末から送信されたデータグラムを受けたデータグラム転送装置が、そのデータグラムのヘッダに記載された転送宛先アドレスに向って前記データグラムをリレーするデータグラム転送システムにおいて、

前記データグラムに関するトラヒック情報に基づいて網に対するインパクトを評価し、その評価値を前記ヘッダに書き込んで、プレファレンス値をつけるトラヒック観測手段を設けたことを特徴とするデータグラム転送システム。

【請求項３】　前記データグラムのヘッダに記載された前記評価値に従った順序によりデータグラムの転送処理を行う手段を有することを特徴とする請求項２に記載のデータグラム転送システム。

【請求項４】　前記データグラム転送装置が、前記データグラムをいずれかの入力インタフェース部からいずれかの出力インタフェース部へ競合することなく転送するバックプレーンスイッチ部を有し、

前記出力インタフェース部が、受信したデータグラムのヘッダから前記プレファレンス値を取得し、そのプレファレンス値の小さいものから順に、輻輳状態とならない範囲で転送すべきデータグラムを選択して、優先的にバッファメモリに書き込むバッファ書き込み制御部を有することを特徴とする請求項２に記載のデ

ータグラム転送システム。

【請求項５】　前記トラヒック情報が、データグラムのサイズまたは連続して送出されるデータグラムの時間間隔であることを特徴とする請求項２に記載のデータグラム転送システム。

【請求項６】　前記トラヒック観測手段が、前記プレファレンス値としてデータグラムのヘッダに存在するデータグラムのサイズフィールドを利用することを特徴とする請求項２に記載のデータグラム転送システム。

【請求項７】　前記トラヒック観測手段が、一つ前の前記データグラムの送出時刻と現在時刻との差の逆数およびこれらの各時刻に送出および到着したデータグラムの大きさに基づいて得られる値をプレファレンス値として計算することを特徴とする請求項２に記載のデータグラム転送システム。

【請求項８】　前記トラヒック観測手段が、前記データグラムのサイズおよび連続するデータグラムの間隔からスライディングウィンドウ方式による平均レートをプレファレンス値として計算することを特徴とする請求項２に記載のデータグラム転送システム。

【請求項９】　前記トラヒック観測手段が、前記データグラムのサイズおよび連続するデータグラムの間隔から観測期間の平均レートをプレファレンス値として計算することを特徴とする請求項２に記載のデータグラム転送システム。

【請求項１０】　前記トラヒック観測手段が、ユーザが送出したデータグラムの個数と受信したデータグラムの個数との差をプレファレンス値として求めることを特徴とする請求項２に記載のデータグラム転送システム。

【請求項１１】　前記バッファ書き込み制御部が、受信したデータグラムのプレファレンス値をプレファレンス値取り出し機能により取得し、このプレファレンス値をキーにしてプレファレンス比較機能にてソーティングを行い、データグラムにプレファレンス値の昇順に順序付けを行い、さらにその順序に従って書き込み制御機能にバッファメモリに対する書き込み処理を行わせることを特徴とする請求項４に記載のデータグラム転送システム。

【請求項１２】　前記バッファメモリが複数の優先順位を持ったクラス別バッファメモリ部を有し、前記バッファ書き込み制御部がプレファレンス値に従っ

てその優先順位を持った複数のクラスを選択し、前記データグラムの書き込みを行うことを特徴とする請求項４に記載のデータグラム転送システム。

【請求項１３】　前記プレファレンス値の取得処理が、データグラムの到着率に応じて変更可能な可変間隔または常に固定の固定間隔の期間ごとに行われることを特徴とする請求項１１に記載のデータグラム転送システム。

【請求項１４】　前記バッファ書き込み制御部では、データグラムの前記バッファメモリへの書き込み前に転送判断を行い、転送を行わないと判断した場合には、前記バッファメモリに空き領域があっても廃棄を行い、転送を行うと判断した場合には前記バッファメモリへの書き込み処理を行うことを特徴とする請求項１１に記載のデータグラム転送システム。

【請求項１５】　前記バッファ書き込み制御部が、既にバッファメモリに書き込んである処理ずみデータグラムのプレファレンス値の総和を計算し、この計算結果に基づく確率計算を行って、その確率に応じてデータグラムの廃棄を行うことを特徴とする請求項１１に記載のデータグラム転送システム。

【請求項１６】　前記バッファ書き込み制御部が、データグラムにプレファレンス値の昇順に順序付けを行った後、その順序付けに従ったバッファメモリの使用量の概算値に基づく確率を求め、得られた確率に応じてデータグラムの廃棄を行うことを特徴とする請求項１１に記載のデータグラム転送システム。

【請求項１７】　前記バッファ書き込み制御部が、適当なタイミングにてスレッショルド値を計算しておき、到着したデータグラムのプレファレンス値と前記スレッショルド値との比較を行い、プレファレンス値がスレッショルド値より大きい場合に前記データグラムを廃棄し、小さい場合に前記バッファメモリへの書き込みを行うことを特徴とする請求項１１に記載のデータグラム転送システム。

【請求項１８】　前記バッファ書き込み制御部が、適当なタイミングにてスレッショルド値を計算しておき、到着したデータグラムのプレファレンス値と前記スレッショルド値との比較を行い、これらの各値の差を入力とする関数により確率を計算し、得られた確率に応じてデータグラムの廃棄を行うことを特徴とする請求項１１に記載のデータグラム転送システム。

【請求項１９】　前記バッファ書き込み制御部が、あらかじめ決められた範囲のデータグラムの到着，転送および廃棄のうちのいずれかの事象が発生した時刻および前記プレファレンス値を保存するプレファレンス値保存機能を有し、これらのプレファレンス値を用いて前記スレッショルド値が計算されることを特徴とする請求項１７または請求項１８に記載のデータグラム転送システム。

【請求項２０】　ユーザによるデータグラムの送出時に、そのデータグラムが網に対してどのようなインパクトを与えるかという評価を行い、この評価を反映したプレファレンス値を計算して、その計算結果を前記データグラムのヘッダに書き込み、前記プレファレンス値による順序付けに従って前記データグラムの転送処理を行うことを特徴とするデータグラム転送方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

この発明は、高速コンピュータ間通信を提供する公衆網におけるベストエフォート型データ通信サービスを提供するためのトラヒック観測装置，データグラム転送システムおよびデータグラム転送方法に関する。

【０００２】

【従来の技術】

インターネットで用いられているインターネットプロトコル（ＩＰ）パケットのように、その配達が網によって保証されないデータユニットをデータグラムと呼ぶ。このようなデータグラムの転送によりデータ通信サービスを実現する網をデータグラム転送網と呼ぶ。このようなデータグラム転送網では、網内のデータグラム転送装置（例えば、インターネットではルータと呼ばれる装置が相当する）が、任意の大きさを持ったデータグラムをそのヘッダに記載された宛先アドレスに向かってリレーすることにより、データ通信サービスが実現される。また、このデータグラム転送装置では、自らの処理能力を超えるデータグラムがごく短い期間に集中して到着した場合（これを輻輳状態と呼ぶ）、到着したデータグラムを適宜廃棄してしまう。このため、データグラム転送網では、「網は個々のデータグラムをその宛先への配達を目指し、網内の転送装置の能力の可能な限りの

転送を行う」というベストエフォート型データ通信サービスを提供することとなる。

【０００３】

図１２は、これまで述べたようなデータグラム転送システムを構成するデータグラム転送装置の一般的な構成を示すブロック図であり、このデータグラム転送装置は、複数の入力インタフェース部（以下、入力Ｉ／Ｆ部という）１と複数の出力インターフェース部（以下、出力Ｉ／Ｆ部という）２とが、一つのバックプレーンスイッチ部３に接続された構成をとっている。このバックプレーンスイッチ部３はいずれか任意の入力Ｉ／Ｆ部１からいずれか任意の出力Ｉ／Ｆ部２へ、内部で競合することなくデータグラムを転送する能力を持っている。

【０００４】

図１３は、図１２における入力Ｉ／Ｆ部１を詳細に示す。この入力Ｉ／Ｆ部１は、ラインＩ／Ｆ部１ａ，転送宛先テーブル１ｂ，転送処理部１ｃおよびデータグラム転送部１ｄから構成されている。そして、この入力Ｉ／Ｆ部１では、ラインＩ／Ｆ部１ａにおいて入力リンクからのデータグラムが受信されると、転送処理部１ｃは転送宛先テーブル１ｂを参照し、宛先アドレスから所望の出力Ｉ／Ｆ部１２を決定するとともに、前記バックプレーンスイッチ部３を介して受信したデータグラムを所望の出力Ｉ／Ｆ部２へ転送させるために、適当なデータグラム転送部１ｄへデータグラムを送る。

【０００５】

図１４は、図１２における出力Ｉ／Ｆ部２を詳細に示す。この出力Ｉ／Ｆ部２は、データグラム受信部２ａ，バッファメモリ２ｂ，バッファ書き込み制御部２ｃ，バッファ読み出し制御部２ｄおよびラインＩ／Ｆ部２ｅから構成されている。そして、この出力Ｉ／Ｆ部２では、バックプレーンスイッチ部３よりデータグラム受信部２ａに受信されたデータグラムは、バッファ書き込み制御部２ｃの処理方法に従って、バッファメモリ２ｂに書き込まれ、出力リンクが利用可能になるのを待つ。すなわち、図１５のフローチャートに示すように、あらかじめ決められた順（通常はＩ／Ｆ番号の昇順など）に入力Ｉ／Ｆ部１を選択し（ステップＳ１）、そこから到着しているデータグラムの有無を調べる（ステップＳ２）。

到着している場合には、これをバッファメモリ２ｂの空き領域に書き込み（ステップＳ３）、未処理の入力Ｉ／Ｆ部１があれば（ステップＳ４）、ステップＳ１以下の処理を繰り返し、なければ終了する。これらの入力Ｉ／Ｆ部１に対する処理は、出力リンクにおける処理単位時間（通常は一つのデータグラムを送出する時間）に対して入力Ｉ／Ｆ部１の数がｎの場合、ｎ倍の速度で行う必要がある。一方、バッファ読み出し制御部２ｄは、ラインＩ／Ｆ２ｅが空きになると、書き込まれた順にバッファメモリ２ｂからラインＩ／Ｆ部２ｅへデータグラムを転送する。そして、このデータグラムはラインＩ／Ｆ部２ｅにより出力リンクへ送信される。

【０００６】

従来のデータグラム転送装置では、この出力Ｉ／Ｆ部２におけるバッファ書き込み制御部２ｄの処理方法として、バッファメモリ２ｂに空きがある限り、到着した順にデータグラムを書き込んで行くという方法が採用されている。この処理方法とバッファ読み出し制御部２ｄにて書き込まれた順に読み出す方法とを組み合わせた方法は、最初に到着したデータグラムが最初に出力されることから、ファーストインファーストアウト（ＦＩＦＯ）方式と呼ばれる。

【０００７】

また、バッファメモリ２ｂに空きがない期間にデータグラムが到着した場合、通常それらはすべて廃棄されるが、特にインターネットにおいては、バッファメモリ２ｂを使い尽くす状況を回避し、多くのデータグラムが一度に廃棄されるという状況を防ぐことで、リンクの利用効率を高めることが可能となる。そのため、インターネットで用いられるルータでは、バッファ書き込み方法として、バッファメモリにたとえ空きがあったとしても、バッファメモリの使用量に応じた適当な確率であらかじめデータグラムを廃棄してしまうというランダムアーリーディテクション（ＲＥＤ）方式が採用されている。

【０００８】

また、図１６はこのＲＥＤ方式による書き込み処理手順を示す。これは、あらかじめ決められた順に入力Ｉ／Ｆ部１を選択し（ステップＳ１）、そこから到着しているデータグラムの有無を調べ（ステップＳ２）、バッファメモリ２ｂの使

用量の概算値を求め（ステップＳ５）、この概算値に基づいた確率を求め（ステップＳ６）、続いて転送判断を行い（ステップＳ７）、この判断結果に基づき、転送を行わない場合にはバッファメモリ２ｂに空きがあっても廃棄を行うという手順をとる。

【０００９】

このように、従来のデータグラム転送装置にあっては、バッファメモリ２ｂへの書き込み方法に注目すると、ＦＩＦＯ方式の場合はデータグラムの到着順に、一方、ＲＥＤ方式の場合でも、バッファメモリ２ｂの使用量に応じた適当な確率でパケットを廃棄するものの、基本的には到着順にバッファメモリ２ｂへデータグラムを書き込んでいる。そして、書き込み時にバッファメモリ２ｂに空きがない場合にはデータグラムを廃棄するというものであり、また、データグラムを転送するか廃棄するかの判断を、データグラムそのものに関する情報ではなく、出力Ｉ／Ｆ部２での情報、つまりバッファメモリ２ｂの利用状況、例えば、バッファメモリに空きがない、バッファメモリの空きが１０％を下回っている等に基づいて行っている。

【００１０】

【発明が解決しようとする課題】

しかしながら、このような方法を用いる従来のデータグラム転送システムにおいては、データグラム転送装置がデータグラムの転送処理を出力Ｉ／Ｆ部２で得られる情報のみしか考慮に入れずデータグラムの到着順に行っており、短い期間に他のユーザより多くのデータグラムを送信することで、送信したデータグラムの廃棄も多くなるが、全体としては他のユーザより多くの網資源を獲得するということが可能となる。多くのユーザがより多くの網資源を獲得することを目的として、このように多くのデータグラムを送出した場合、網内のそれぞれのデータグラム転送装置は輻輳状態に陥り、送出されたデータグラムの多くを廃棄することとなる。この結果、廃棄されずに最終的な宛先まで到着するデータグラムの数が急激に減少し、網全体としての実効的なデータ転送能力が低下した状態、つまり輻輳崩壊状態に陥ってしまうという課題があった。

【００１１】

この発明は前記課題を解決するものであり、データグラム転送装置の出力Ｉ／Ｆ部におけるバッファ書き込み制御部が、データグラムの送出に対する網の評価により得られる情報（プレファレンス値）を利用し、データグラムを到着順以外の方法で転送処理を行うことで、ユーザがより多くの網資源を獲得しようとして必要以上のデータグラムを送出することを抑制でき、これにより輻輳崩壊状態に陥らない安定したデータグラム通信網を実現できるデータグラム転送システムを得ることを目的とする。

【００１２】

【課題を解決するための手段】

前記目的達成のため、請求項１の発明にかかるトラヒック観測装置は、現在よりあらかじめ決められた期間前までの、ユーザが網内に送信または網から受信したデータグラムに関するトラヒック情報を観測するトラヒック観測機能と、該トラヒック観測機能により得られたトラヒック情報からデータグラムの送出に対する網の評価により得られるプレファレンス値を計算するプレファレンス値計算機能とを有し、プレファレンス値挿入機能に、前記プレファレンス値計算機能により計算されたプレファレンス値を前記評価値として前記送信されるデータグラムのヘッダに書き込ませるようにしたものである。

【００１３】

また、請求項２の発明にかかるデータグラム転送システムは、ユーザ端末から送信されたデータグラムを受けたデータグラム転送装置が、そのデータグラムのヘッダに記載された転送宛先アドレスに向って前記データグラムをリレーするデータグラム転送システムに、前記データグラムに関するトラヒック情報に基づいてユーザのトラヒックの網に対するインパクトを評価し、その評価値を前記ヘッダに書き込んで、この評価に従った順序による転送処理を行わせるトラヒック観測手段を設けたものである。

【００１４】

また、請求項３の発明ないかかるデータグラム転送システムは、前記データグラムのヘッダに記載された前記評価値に従った順序によりデータグラムの転送処理を行う手段を設けたものである。

【００１５】

また、請求項４の発明にかかるデータグラム転送システムは、前記データグラム転送装置に、前記データグラムをいずれかの入力インタフェース部からいずれかの出力インタフェース部へ競合することなく転送するバックプレーンスイッチ部を設け、前記出力インタフェース部が、受信したデータグラムのヘッダから前記プレファレンス値を取得し、そのプレファレンス値の小さいものから順に、輻輳状態とならない範囲で転送すべきデータグラムを選択して、優先的にバッファメモリに書き込むバッファ書き込み制御部を設けたものである。

【００１６】

また、請求項５の発明にかかるデータグラム転送システムは、前記トラヒック情報を、データグラムのサイズまたは連続して送出されるデータグラムの時間間隔としたものである。

【００１７】

また、請求項６の発明にかかるデータグラム転送システムは、前記トラヒック観測手段に、前記プレファレンス値としてデータグラムのヘッダに存在するデータグラムのサイズフィールドを利用させるようにしたものである。

【００１８】

また、請求項７の発明にかかるデータグラム転送システムは、前記トラヒック観測手段に、一つ前の前記データグラムの送出時刻と現在時刻との差の逆数をプレファレンス値として計算させるようにしたものである。

【００１９】

また、請求項８の発明にかかるデータグラム転送システムは、前記トラヒック観測手段に、前記データグラムのサイズおよび連続するデータグラムの間隔からスライディングウィンドウ方式による平均レートをプレファレンス値として計算させるようにしたものである。

【００２０】

また、請求項９の発明にかかるデータグラム転送システムは、前記トラヒック観測手段に、前記データグラムのサイズおよび連続するデータグラムり間隔から観測期間の平均レートをプレファレンス値として計算させるようにしたものであ

る。

【００２１】

また、請求項１０の発明にかかるデータグラム転送システムは、前記トラヒック観測手段に、ユーザが送出したデータグラムの個数と受信したデータグラムの個数との差をプレファレンス値として求めさせるようにしたものである。

【００２２】

また、請求項１１の発明にかかるデータグラム転送システムは、前記バッファ書き込み制御部で、受信したデータグラムのプレファレンス値をプレファレンス値取り出し機能により取得し、このプレファレンス値をキーにしてプレファレンス比較機能にてソーティングを行い、データグラムにプレファレンス値の昇順に順序付けを行い、さらにその順序に従って書き込み制御機能にバッファメモリに対する書き込み処理を行わせるようにしたものである。

【００２３】

また、請求項１２の発明にかかるデータグラム転送システムは、前記バッファメモリに複数の優先順位を有するクラス別バッファメモリ部を持たせ、前記バッファ書き込み制御部にプレファレンス値に従ってその優先順位を持った複数のクラスを選択し、前記データグラムの書き込みを行わせるようにしたものである。

【００２４】

また、請求項１３の発明にかかるデータグラム転送システムは、前記プレファレンス値の取得処理を、データグラムの到着率に応じて変更可能な可変間隔または常に固定の固定間隔の期間ごとに行うようにしたものである。

【００２５】

また、請求項１４の発明にかかるデータグラム転送システムは、前記バッファ書き込み制御部で、データグラムの前記バッファメモリへの書き込み前に転送判断を行わせ、転送を行わないと判断した場合には、前記バッファメモリに空き領域があっても廃棄を行わせ、転送を行うと判断した場合には前記バッファメモリへの書き込み処理を行わせるようにしたものである。

【００２６】

また、請求項１５の発明にかかるデータグラム転送システムは、前記バッファ

書き込み制御部で、既にバッファメモリに書き込んである処理ずみデータグラムのプレファレンス値の総和を計算させ、この計算結果に基づく確率計算を行わせて、その確率に応じてデータグラムの廃棄を行わせるようにしたものである。

【００２７】

また、請求項１６の発明にかかるデータグラム転送システムは、前記バッファ書き込み制御部で、データグラムにプレファレンス値の昇順に順序付けを行わせた後、その順序付けに従ったバッファメモリの使用量の概算値に基づく確率を求めさせ、得られた確率に応じてデータグラムの廃棄を行わせるようにしたものである。

【００２８】

また、請求項１７の発明にかかるデータグラム転送システムは、前記バッファ書き込み制御部で、適当なタイミングにてスレッショルド値を計算しておき、到着したデータグラムのプレファレンス値と前記スレッショルド値との比較を行わせ、プレファレンス値がスレッショルド値より大きい場合に前記データグラムを廃棄させ、小さい場合に前記バッファメモリへの書き込みを行わせるようにしたものである。

【００２９】

また、請求項１８の発明にかかるデータグラム転送システムは、前記バッファ書き込み制御部にプレファレンス値とスレッショルド値の各値の差を入力とする関数により確率を計算させ、得られた確率に応じてデータグラムの廃棄を行わせるようにしたものである。

【００３０】

また、請求項１９の発明にかかるデータグラム転送システムは、前記バッファ書き込み制御部に、あらかじめ決められた範囲のデータグラムの到着，転送および廃棄のうちのいずれかの事象が発生した時刻および前記プレファレンス値を保存するプレファレンス値保存機能を持たせ、これらのプレファレンス値を用いて前記スレッショルド値を計算させるようにしたものである。

【００３１】

また、請求項２０の発明にかかるデータグラムの転送方法は、ユーザによるデ

ータグラムの送出時に、そのデータグラムが網に対してどのようなインパクトを与えるかという評価を行い、この評価を反映したプレファレンス値を計算して、その計算結果を前記データグラムのヘッダに書き込み、前記プレファレンス値による順序付けに従って前記データグラムの転送処理を行うようにしたものである。

【００３２】

【発明の実施の形態】

以下、この発明の実施の一形態を図について説明する。この発明のデータグラム転送システムで実行されるプレファレンス値を用いた選択的データグラム転送方法は、以下の２点を組み合わせて実行される。

１．網はユーザのデータグラム送出時にそのデータグラムが網に対しどのようなインパクト（負荷）を与えるかという評価を行い、それを反映させた値としてプレファレンス値を計算し、それをデータグラムのヘッダに付与する。ここで、網に対してのインパクトが少ないものに対して、小さいプレファレンス値を与える。

２．網内のデータグラム転送装置においては、データグラムの到着順ではなく、このプレファレンス値による順序付けに従った順にデータグラムの転送処理を行う。具体的には網に対してのインパクトが少ないことを意味するプレファレンス値の小さいものほど高い確率で優先して転送を行うよう制御する。

【００３３】

図１３は前記データグラムの転送方法を実行するデータグラム転送システムを示すブロック図であり、ここでは一般的な網とユーザ端末との接続形態の様子とこの発明を実現する装置の一つであるトラヒック観測装置が導入される位置を具体的に示してある。これによればユーザ端末４からライン終端装置５およびアクセス網６を介して送信されたデータグラムは、網内のデータグラム転送装置Ｄにおいて、図１に示すような入力Ｉ／Ｆ部１の転送宛先テーブル１ｂに従って、宛先アドレスの示す端末までリレーされて行く。このアクセス網を終端するライン終端装置７と網内の最初のデータグラム転送装置Ｄとの間にトラヒック観測手段としてのトラヒック観測装置８が接続されている。

【００３４】

図２はこのトラヒック観測装置８の構成を示し、これがトラヒック観測機能８ａ，プレファレンス値計算機能８ｂおよびプレファレンス値挿入機能８ｃにより構成され、これらのうちトラヒック観測機能８ａは、現在よりあらかじめ決められた期間前までの、ユーザ網内に送信、あるいは網より受信したデータグラムに関するトラヒック情報、例えばデータグラムのサイズや連続して送出されたデータグラムの時間間隔などを観測し、保持する機能である。また、プレファレンス値計算機能８ｂは、トラヒック観測機能８ａにより保持されている情報から決められた計算式を用いて、プレファレンス値を計算する機能である。また、プレファレンス値挿入機能８ｃは、到着したデータグラムに、プレファレンス値計算機能８ｂにより計算された値をデータグラムのヘッダに書き込む機能である。すなわち、このようなトラヒック観測装置８では、トラヒック観測機能８ａにより得られたデータを用いて、プレファレンス値計算機能８ｂにより、ユーザのトラヒックの網に対するインパクト（負荷）を評価してその数値化を行い、それをデータグラムのヘッダに書き込むことを実現する。そして、前記トラヒック観測装置８のプレファレンス値挿入機能８ｃにより書き込まれるプレファレンス値は、図３に示すように、転送されるデータグラムのヘッダ部分におけるプレファレンス値を書き込むフィールドに保存する。このフィールドの値は、データグラム転送網内を転送される間変更されず、各データグラム転送装置Ｄに、網がトラヒック観測装置８において行った評価結果を伝える機能を実現する。

【００３５】

図４は、この発明のデータグラム転送装置Ｄの出力Ｉ／Ｆ部２におけるバッファ書き込み制御部２ｃの構成を示し、これがプレファレンス値取り出し機能２ｃ１、プレファレンス値比較機能２ｃ２および書き込み制御機能２ｃ３から構成される。これらのうち、プレファレンス値取り出し機能２ｃ１は、到着したデータグラムのヘッダよりプレファレンス値を取り出して保持し、プレファレンス値比較機能２ｃ２は、保持されている各プレファレンス値の比較を行い、到着したデータグラムに対しプレファレンス値のより小さいものが先になるような順序を与える。さらに、書き込み制御機能２ｃ３は、プレファレンス値比較機能２ｃ２の

結果を用い、プレファレンス値の小さいものから順に、輻輳状態とならない範囲で転送するデータグラムを選択し、それらを図３に示すようなバッファメモリ２ｂへ書き込む制御を行い、また、選択されなかったデータグラムは廃棄する制御を行う。

【００３６】

図５は、この発明における書き込み制御機能２ｃ３による書き込み処理方法を示すフローチャートである。これは、従来のＦＩＦＯ方式とは違い、最初にプレファレンス値取り出し機能２ｃ１により到着しているデータグラムのプレファレンス値のリストを取得する（ステップＳ１１）。次に、プレファレンス値比較機能２ｃ２により、この値をキーにしソーティングを行い（ステップＳ１２）、データグラムにプレファレンス値の昇順に順序付けを行う（ステップＳ１３）。そして、その順序に従って、書き込み制御機能２ｃ３によりバッファメモリ２ｂへの書き込み処理を行い（ステップＳ１４）、未処理のデータグラムがある場合には（ステップＳ１５）、ステップＳ１３以下の処理を実行し、ない場合には処理を終了する。

【００３７】

この発明では、データグラムはデータグラム転送装置Ｄにおいて、到着順に処理されるのではなく、ヘッダに存在するプレファレンス値の小さいものが優先的に転送されることとなる。このプレファレンス値は、網がユーザのトラヒックを観測することで付与するため、たとえ、あるユーザが短い期間に多くのデータグラムを送信したとしても、それらのデータグラムにはトラヒック観測装置が大きなプレファレンス値をつけるようになり、データグラム転送装置Ｄでより低いプレファレンスをもつデータグラムよりも高い確率で廃棄されてしまう。従って、短い期間に多くのデータグラムを送信することでデータグラムの廃棄も多くなるが、結果的により多くの網資源を獲得することが不可能となる。その結果、ユーザが多くの網資源の獲得を目的として不必要なデータグラムを送出することを抑制でき、データグラム転送網が輻輳崩壊状態に陥る危険性を低くすることが可能となる。

【００３８】

前記のプレファレンス値を用いた選択的データグラム転送方法は、網によるユーザトラヒックの評価とその数値化およびその数値による順序付けによる優先転送処理の二つの機能的要素からなり、これらの要素は、それぞれトラヒック観測装置８におけるプレファレンス値計算機能８ｂおよびデータグラム転送装置Ｄにおけるバッファ書き込み制御部２ｃの処理方法として実施されるものである。以下では、これらの二つについての実施例をそれぞれ述べる。

【００３９】

前記のように、バッファ書き込み制御部２ｃでは、最初にプレファレンス値取り出し機能２ｃ１により到着しているデータグラムのプレファレンス値のリストを取得し、プレファレンス比較機能として、このプレファレンス値をキーにしてソーティングを行い、データグラムにプレファレンス値の昇順に順序付けを行い、さらにその順序に従って、バッファメモリ２ｂへの書き込み処理を行っている。そしてプレファレンス値の取得処理は、適当な期間Ｔ毎に行う。この期間Ｔを長くすれば、その間に到着してソーティングの対象となるデータグラムの数が増加するため、多くのデータグラム間で優先順位の比較ができ、この発明の効果は大きくなるが、到着してから転送処理が行われ、出力リンクへ送信されるまでの遅延時間が大きくなる。この期間Ｔは、その点を考慮して最適な値を選択する。また、常に固定した期間でもよいが、データグラムの到着率に応じて適応的に期間を変更することも可能である。

【００４０】

また、図６に示すように、図５の処理手順に加えてデータグラムのバッファメモリ２ｂへの書き込みをする前に、転送判断を別途行い（ステップＳ１６）、この判断結果に基づき、転送を行わない場合には（ステップＳ１７）、バッファメモリ２ｂに空き領域があっても廃棄を行うという処理手順をとることも可能である。一つのより具体的な例を図７に示す。既に処理を行いバッファメモリに書き込みを行ったデータグラムのプレファレンス値の和を計算し（ステップＳ１８）、それに基づいた確率を求め（ステップＳ１９）、その確率に応じて廃棄を行うことが可能となる。この場合、後に処理されるデータグラム、つまりプレファレンス値の大きいデータグラムほど高い確率で廃棄されることになり、この処理を

行わない図５の場合に比べてこの発明の効果は大きくなる。また、この発明の書き込み処理方法は、インターネットのルータに用いられている前記ＲＥＤ方式との併用も可能である。図８はこのＲＥＤ方式を併用した書き込み処理手順を示すフローチャートである。これによれば、最初にプレファレンス値のリストを取得し（ステップＳ４０）、ソーティングを行うことで（ステップＳ１２）、得られるデータグラムの順序付けに従って（ステップＳ１３）、ＲＥＤ方式の転送処理手順が行われる。すなわち、バッファメモリ２ｂの使用量の概算を行い（ステップＳ２０）、この概算値に基づいた確率を求め（ステップＳ２１）、続いて転送判断を行った後（ステップＳ１７）、その判断結果に従ってバッファメモリ２ｂへの書き込みまたは廃棄を行うという手順をとる。

【００４１】

ところで、図６～図８の例では、プレファレンス値の取得処理を行う期間Ｔを適当に設定する必要がある。それに対して、図９に示す例では、到着したデータグラムは前記のように同時に到着したものが一度に処理されるのではなく、順番に一つずつ処理され、プレファレンス値を取得した後ソーティングを行うという処理を行わない。すなわち、あらかじめ決められた順に入力Ｉ／Ｆ部１を選択し（ステップＳ３１）、そこから到着しているデータグラムの有無を調べ（ステップＳ３２）、続いて図１０に示すようなプレファレンス値取り出し機能２ｃ１にてプレファレンス値の取得を行って（ステップＳ３３）、このプレファレンス値の保存をプレファレンス値保存機能２ｃ４に行う。さらにスレッショルド計算機能２ｃ５において、適当なタイミングにてスレッショルド値を計算しておき（ステップＳ３４）、到着したデータグラムのプレファレンス値とスレッショルド値との比較を行い（ステップＳ３５）、プレファレンス値がスレッショルド値より大きい場合は、廃棄され、そうでなければ転送を行うという方法で、選択的なバッファ書き込み処理を行う（ステップＳ１４）。この例ではデータグラムの選択順序を変える必要がないため、従来法からの改良が比較的簡単であるという特徴を持つ。なお、プレファレンス値とスレッショルド値の比較を行う場合に、プレファレンス値とスレッショルド値の差を入力とする関数により確率を計算し（ステップＳ３５）、その確率によりデータグラムを廃棄する（ステップＳ１７）と

いう方法も採用可能である。

【００４２】

この発明におけるデータグラム転送装置Ｄにおけるバッファ書き込み制御部２ｃは、あらかじめ決められた範囲のデータグラムの到着，転送，廃棄等のいずれかの事象が起った時刻と、そのプレファレンス値を保存する機能を具備している。また、前記スレッショルド値は、それらの各値を用いて以下に示すような値として計算する。

１．転送した過去ｎ個のパケットに関する平均値／メジアン／平均＋分散／平均＋標準偏差／移動平均／分布の最頻値

２．到着した過去ｎ個のパケットに関する平均値／メジアン／平均＋分散／平均＋標準偏差／移動平均／分布の最頻値

３．廃棄された過去ｎ個のパケットに関する平均値／メジアン／平均＋分散／平均＋標準偏差／移動平均／分布の最頻値

４．過去ｔ秒以内に転送したパケットに関する平均値／メジアン／平均＋分散／平均＋標準偏差／移動平均／分布の最頻値

５．過去ｔ秒以内に到着したパケットに関する平均値／メジアン／平均＋分散／平均＋標準偏差／移動平均／分布の最頻値

６．過去ｔ秒以内に廃棄したパケットに関する平均値／メジアン／平均＋分散／平均＋標準偏差／移動平均／分布の最頻値

７．過去ｔ秒以内に転送した、過去ｎ個のパケットに関する平均値／メジアン／平均＋分散／平均＋標準偏差／移動平均／分布の最頻値

８．過去ｔ秒以内に到着した、過去ｎ個のパケットに関する平均値／メジアン／平均＋分散／平均＋標準偏差／移動平均／分布の最頻値

９．過去ｔ秒以内に廃棄した、過去ｎ個のパケットに関する平均値／メジアン／平均＋分散／平均＋標準偏差／移動平均／分布の最頻値

【００４３】

次に、出力Ｉ／Ｆ部２に複数のクラス別バッファメモリ部を持つバッファメモリ２ｂを図１１に示す。これらのクラス別バッファメモリ部はそれぞれ優先順位をもっており、その優先順位に従ってバッファ読み出し制御部２ｄにより読み出

しが行われる。バッファ書き込み制御部２ｃは、プレファレンス値比較機能により与えられた順位に従って、複数の優先順位を持ったクラス別バッファメモリ部の選択を行い、そこにデータグラムを書き込む。具体的には以下の処理をバッファ書き込み制御部２ｃにおいて行う。すなわち、いま、ｎ個の優先順位を持ったキューが存在し、１～ｎまでの番号が割り振られている場合を考え、番号が小さいものほど優先順位は高いものとする。また、プレファレンス値が取りうる値をｎ個の区間に分割し、プレファレンスの昇順に１～ｎの番号を割り振る。さらに、区間ｉのプレファレンス値を持ったデータグラムをｉのキューに割り振る。以上の処理をバッファ書き込み制御部において行うことにより、小さいプレファレンス値を持つデータグラムほど優先度の高いバッファメモリに割り振られることとなり、この発明の効果を実現することが可能となる。

【００４４】

この発明では、通常のデータグラム（例えばＩＰ）のヘッダに存在するデータグラムのサイズフィールドをプレファレンス値として利用することができる。これに対し、以下に示す例では、トラヒック観測装置８においてトラヒックの観測により計算された値をデータグラムのヘッダの特別なフィールドに挿入することを基本としている。しかし、この例ではデータグラムのフォーマットに特別なフィールドは必要ない、但しデータグラムのサイズは網がつけるものではなくユーザ端末がつける値であるため、トラヒック観測装置８は、書かれているデータグラムのサイズの値と実際のサイズが正しいかどうかというテストを行い、正しくない場合にはそのデータグラムを廃棄する処理を行う。

【００４５】

また、プレファレンス値の他の計算機能として、一つ前のデータグラムの送出時刻と、現在時刻との差の逆数（データグラム転送間隔の逆数）を用いることができる。この場合には、現在時刻をｔ、現在到着したデータグラムの大きさをＭとし、また、一つ前のデータグラムの送出時刻をｔｉ、データグラムの大きさをＭｉ、αを影響度の度合を調整する定数とすると、α（Ｍ／Ｍｉ）｛１／（ｔ－ｔｉ）｝式で表現される値を用いる。この値は、瞬間的なユーザのトラヒックの速度を意味する値であるため、この値が大きいほど網に対するインパクトが大き

いと判断することが可能である。

【００４６】

さらに、プレファレンス値の他の計算機能として、データグラムのサイズおよび連続するデータグラムの間隔から計算する平均レート（スライディングウィンドウ方式）を用いることができる。この場合には、現在からｉ個以前のデータグラムの大きさをＭｉ、到着時刻をｔｉ、遡るデータグラム数をｎ、影響の度合を調整する定数をαｉとすると（ｎ，αｉはあらかじめ設定された値）、平均レートＶＫは、αｉＭｉ／（Ｔｉ－Ｔｉ+1）のｉが０～ｎまでの和として求められる。

【００４７】

この値ＶＫは前記データグラムの転送間隔の逆数を用いた場合より長い時間間隔でみた平均レートとなるので、やはり網に対するインパクトを評価することが可能である。

【００４８】

また、プレファレンス値の他の計算機能として、データグラムのサイズおよび、連続するデータグラムの間隔から計算する観測期間の平均レートを用いることができる。すなわち、観測期間をＴＡとすると、この観測期間ＴＡに到着したデータグラムのサイズの和を、観測期間ＴＡで除算したものを利用する。また、データグラムのサイズの和そのものを用いることも可能であり、また、その和をアクセス網の物理リンク速度より決まる和の最大値で除算し、正規化を行った値とすることも可能である。

【００４９】

また、プレファレンス値の他の計算機能として、ユーザが送出したデータグラムの個数と、受信したデータグラムの個数の差を用いることができる。通常、データグラム転送網を適正に用いる場合には、データグラムの配達が保証されないため、ユーザ端末は正しく受信できたデータグラムに対し、正しく受信できたことを送信ユーザ端末に知らせるＡＣＫデータパケットを送信する。従って、このような適正な方法を用いてデータグラム通信網を利用する場合には、送信したデータグラム数と受信したデータグラム数はほぼ等しくなる。一方、この発明が課

題とするような方法をユーザがとる場合には、網内で多くのデータグラムが廃棄され、それに対するＡＣＫデータパケットが送信されないため、送信データグラム数と受信データグラム数には大きな開きができると考えられる。従って、この差をプレファレンス値とすれば、適正な利用方法を行っているものがより優先的に転送されることとなり、この発明の効果を得ることが可能となる。

【００５０】

【発明の効果】

以上のように、この発明によれば、データグラム転送／廃棄方法を実現するデータグラム転送装置を用いることで、ユーザが多くの網資源を獲得しようとして必要以上のデータグラムを送出することを抑制でき、これによって輻輳崩壊状態に陥らない安定したデータグラム通信網を実現できるという効果が得られる。

【図面の簡単な説明】

【図１】　この発明の実施の一形態によるデータグラム転送システムを示すブロック図である。

【図２】　図１におけるトラヒック観測装置の詳細を示すブロック図である。

【図３】　この発明におけるデータグラムのヘッダフィールドを示す説明図である。

【図４】　この発明における出力インタフェース部内のバッファ書き込み制御部を示すブロック図である。

【図５】　この発明における書き込み制御機能によるバッファメモリへの書き込み手順を示すフローチャートである。

【図６】　この発明における書き込み制御機能によるバッファメモリへの他の書き込み手順を示すフローチャートである。

【図７】　この発明における書き込み制御機能によるバッファメモリへの他の書き込み手順を示すフローチャートである。

【図８】　この発明における書き込み制御機能によるバッファメモリへの他の書き込み手順を示すフローチャートである。

【図９】　この発明における書き込み制御機能によるバッファメモリへの他

の書き込み手順を示すフローチャートである。

【図１０】　この発明における出力インタフェース部内のバッファ書き込み制御部の他の例を示すブロック図である。

【図１１】　この発明における出力インタフェース部の他の例を示すブロック図である。

【図１２】　従来およびこの発明のデータグラム転送装置の基本構成を示すブロック図である。

【図１３】　図１２における入力インタフェース部の従来例を示すブロック図である。

【図１４】　図１２における出力インタフェース部の従来例を示すブロック図である。

【図１５】　従来のファーストインファーストアウト方式によるデータグラム転送手順を示すフローチャートである。

【図１６】　従来のランダムアーリーディテクション方式によるデータグラム転送手順を示すフローチャートである。

【符号の説明】

１　入力Ｉ／Ｆ部（入力インタフェース部）

２　出力Ｉ／Ｆ部（出力インタフェース部）

２ｃ　バッファ書き込み制御部

２ｃ４　　プレファレンス値保存機能

３　バックプレーンスイッチ部

８　トラヒック観測装置

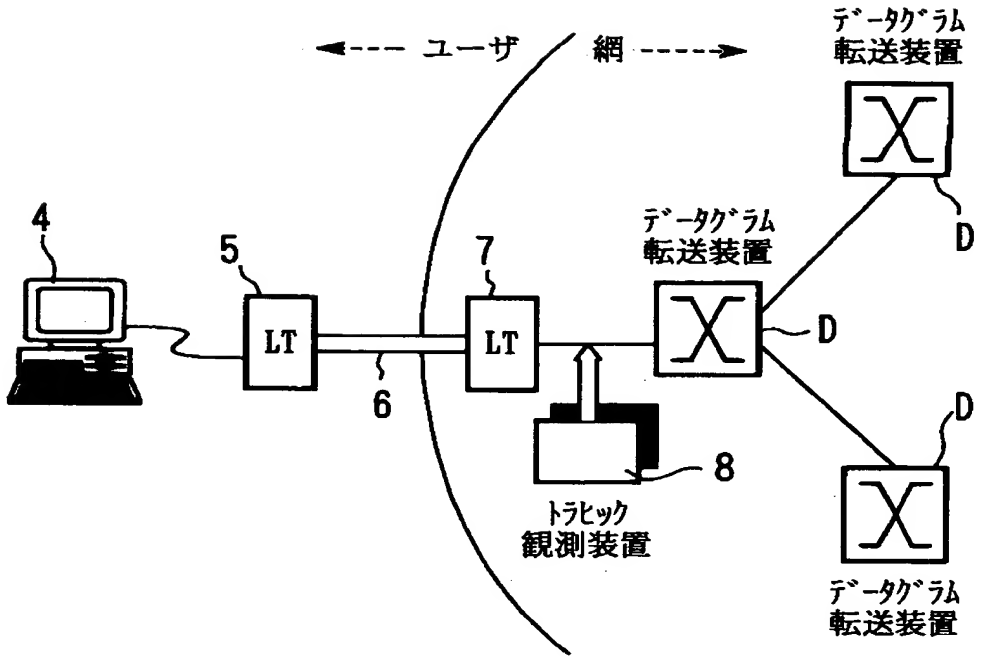
８ａ　トラヒック観測機能

８ｂ　プレファレンス値計算機能

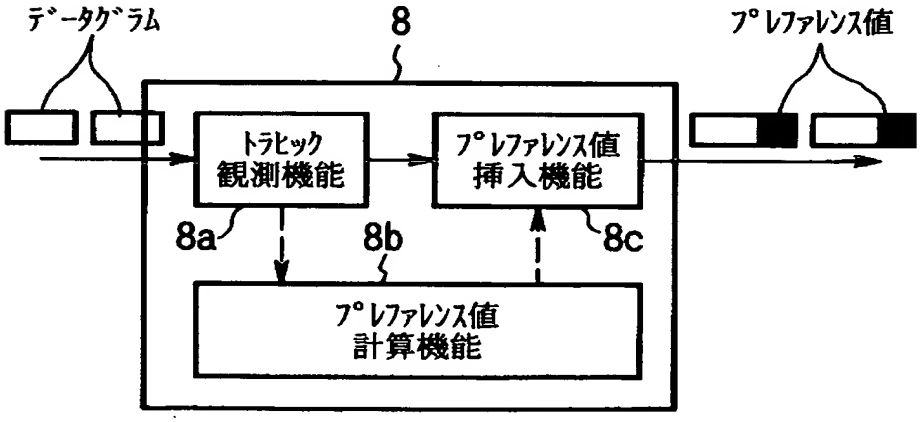
８ｃ　プレファレンス値挿入機能

【書類名】　　図面

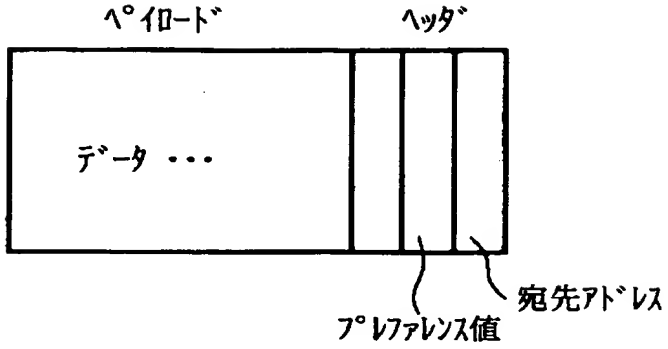
【図１】


ユーザ
網
ﾃﾞｰﾀｸﾞﾗﾑ
転送装置
4
5
7
LT
LT
6
ﾃﾞｰﾀｸﾞﾗﾑ
転送装置
D
D
D
8
ﾄﾗﾋｯｸ
観測装置
ﾃﾞｰﾀｸﾞﾗﾑ
転送装置

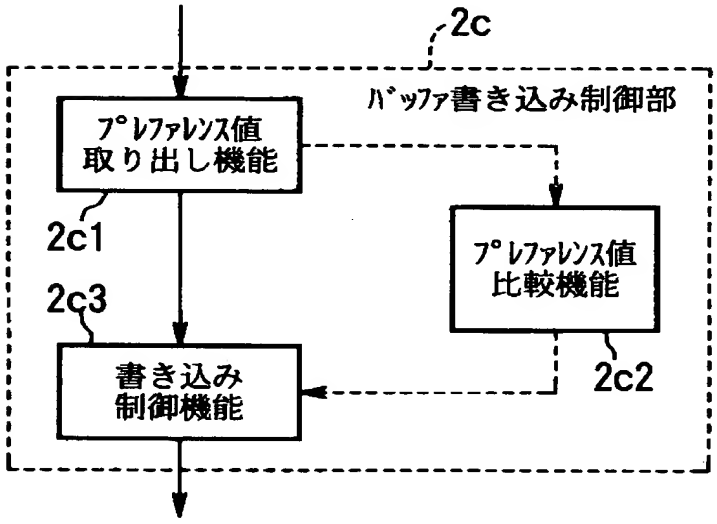
【図２】


ﾃﾞｰﾀｸﾞﾗﾑ
8
ﾌﾟﾚﾌｧﾚﾝｽ値
ﾄﾗﾋｯｸ
観測機能
ﾌﾟﾚﾌｧﾚﾝｽ値
挿入機能
8a
8b
8c
ﾌﾟﾚﾌｧﾚﾝｽ値
計算機能

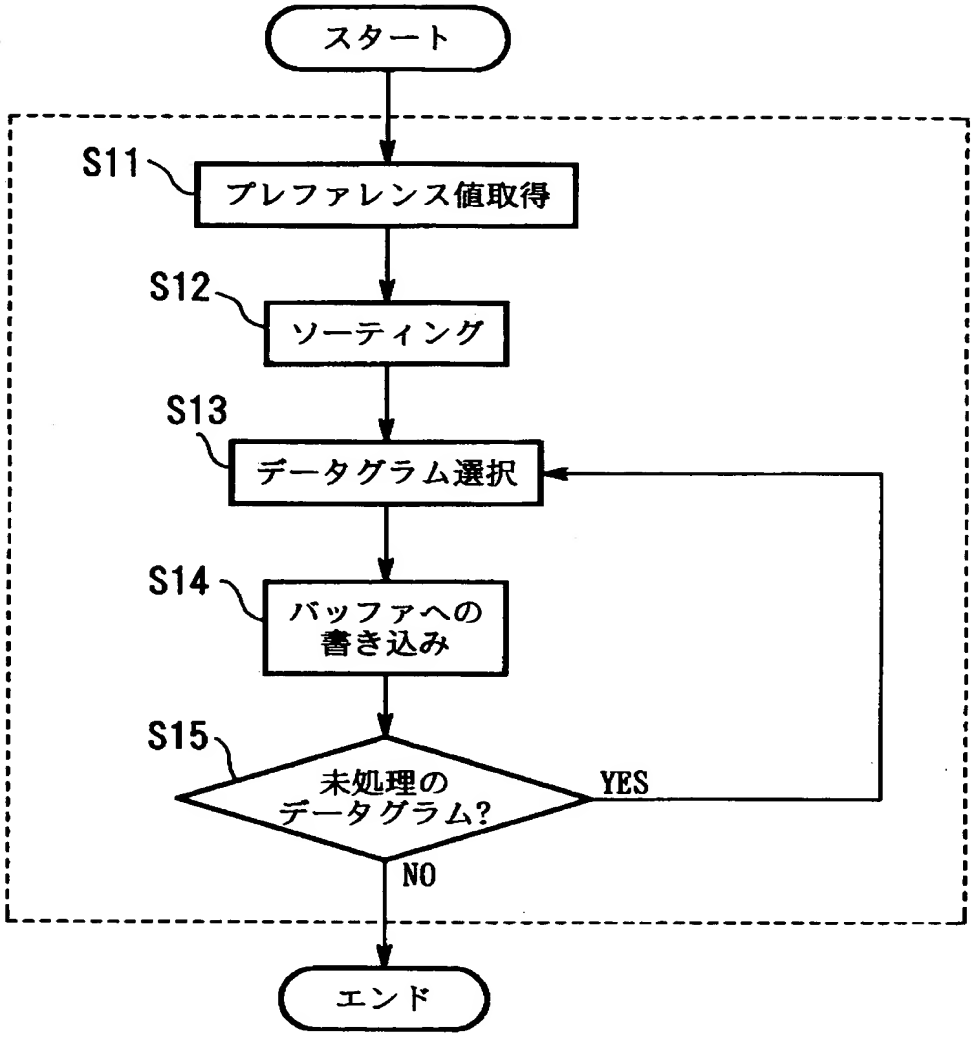
【図３】


ﾍﾟｲﾛｰﾄﾞ
ﾍｯﾀﾞ
ﾃﾞｰﾀ ・・・
宛先ｱﾄﾞﾚｽ
ﾌﾟﾚﾌｧﾚﾝｽ値

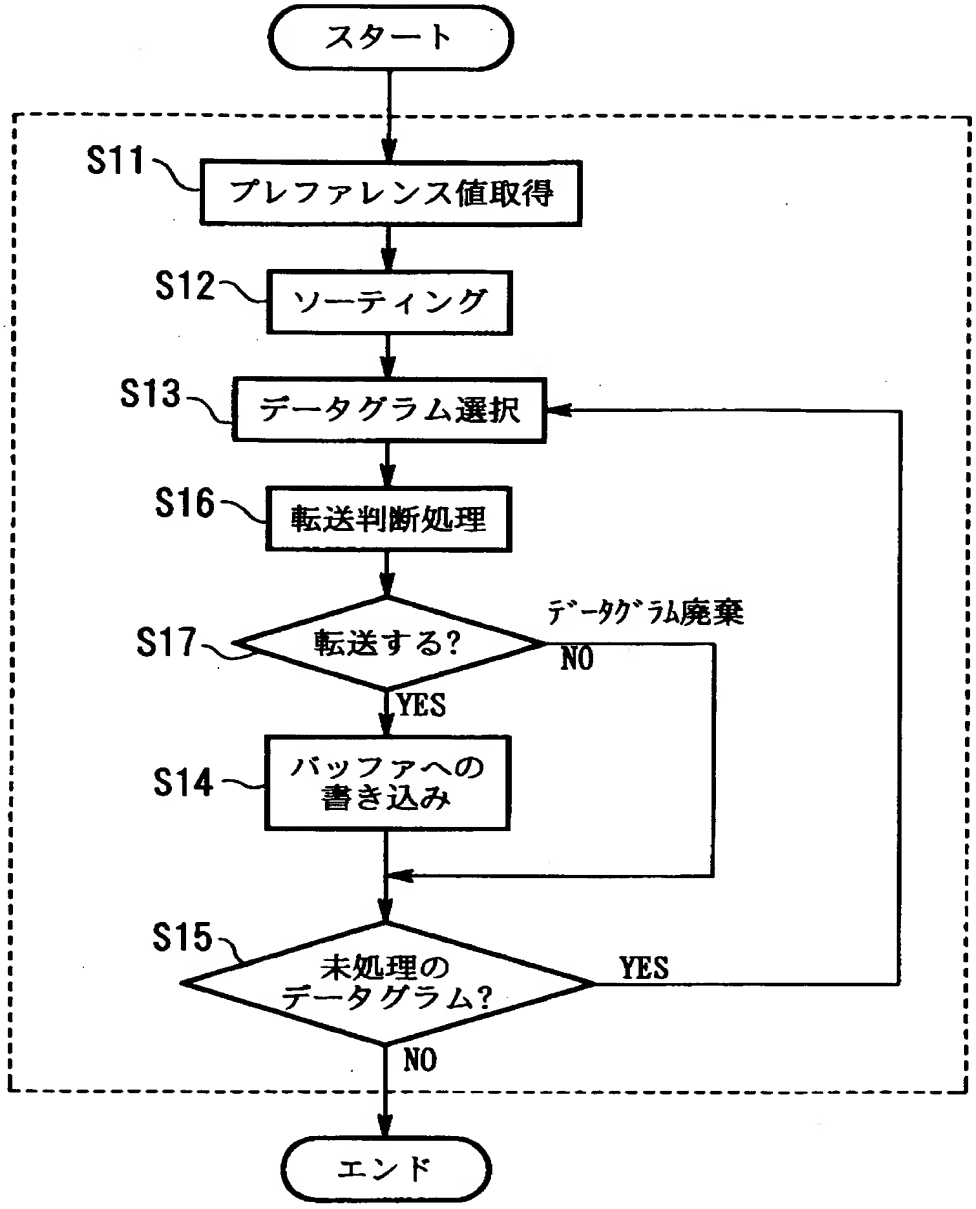
【図４】


2c
ﾊﾞｯﾌｧ書き込み制御部
ﾌﾟﾚﾌｧﾚﾝｽ値
取り出し機能
2c1
2c3
ﾌﾟﾚﾌｧﾚﾝｽ値
比較機能
2c2
書き込み
制御機能

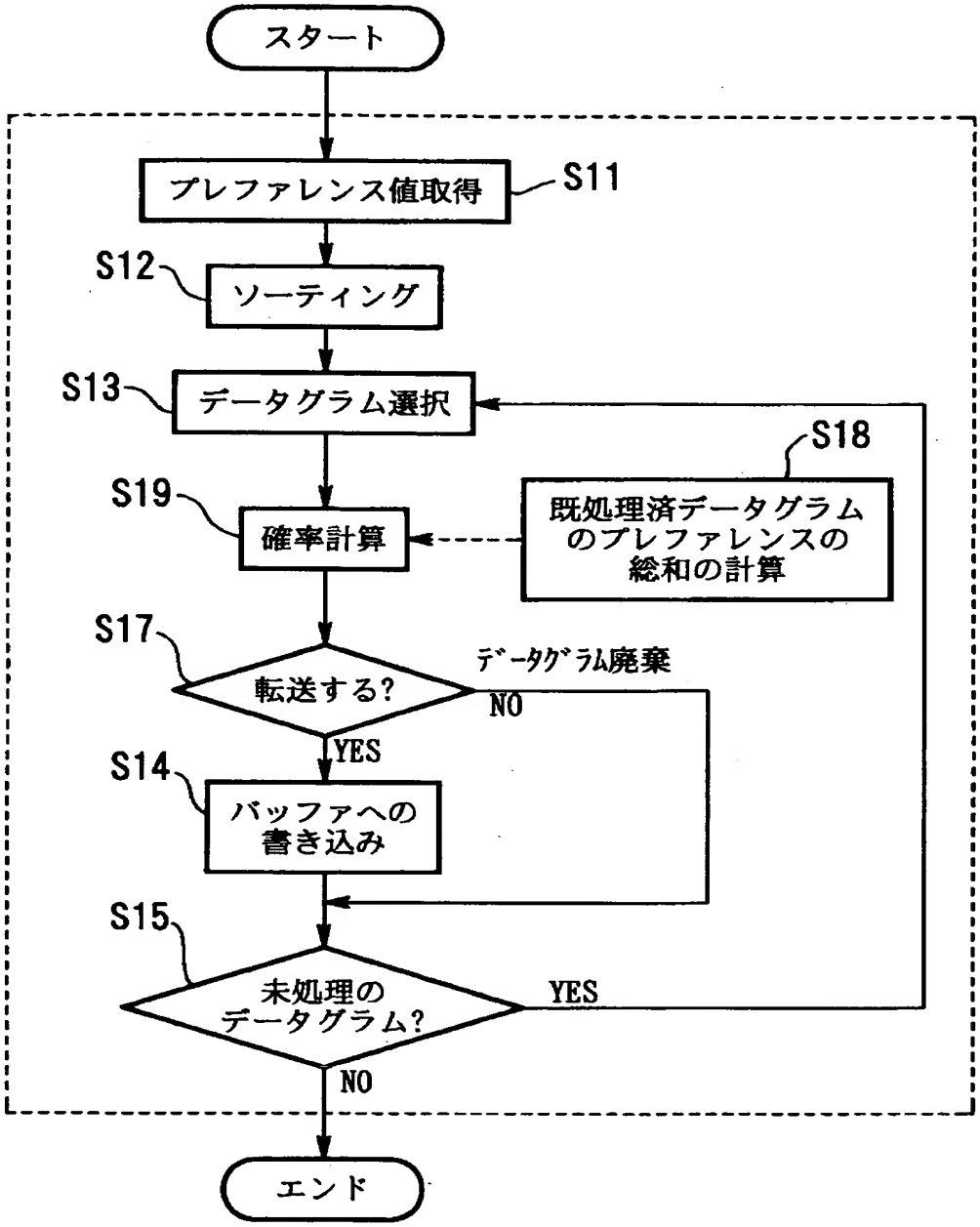
【図５】


スタート
S11
プレファレンス値取得
S12
ソーティング
S13
データグラム選択
S14
バッファへの
書き込み
S15
未処理の
データグラム?
YES
NO
エンド

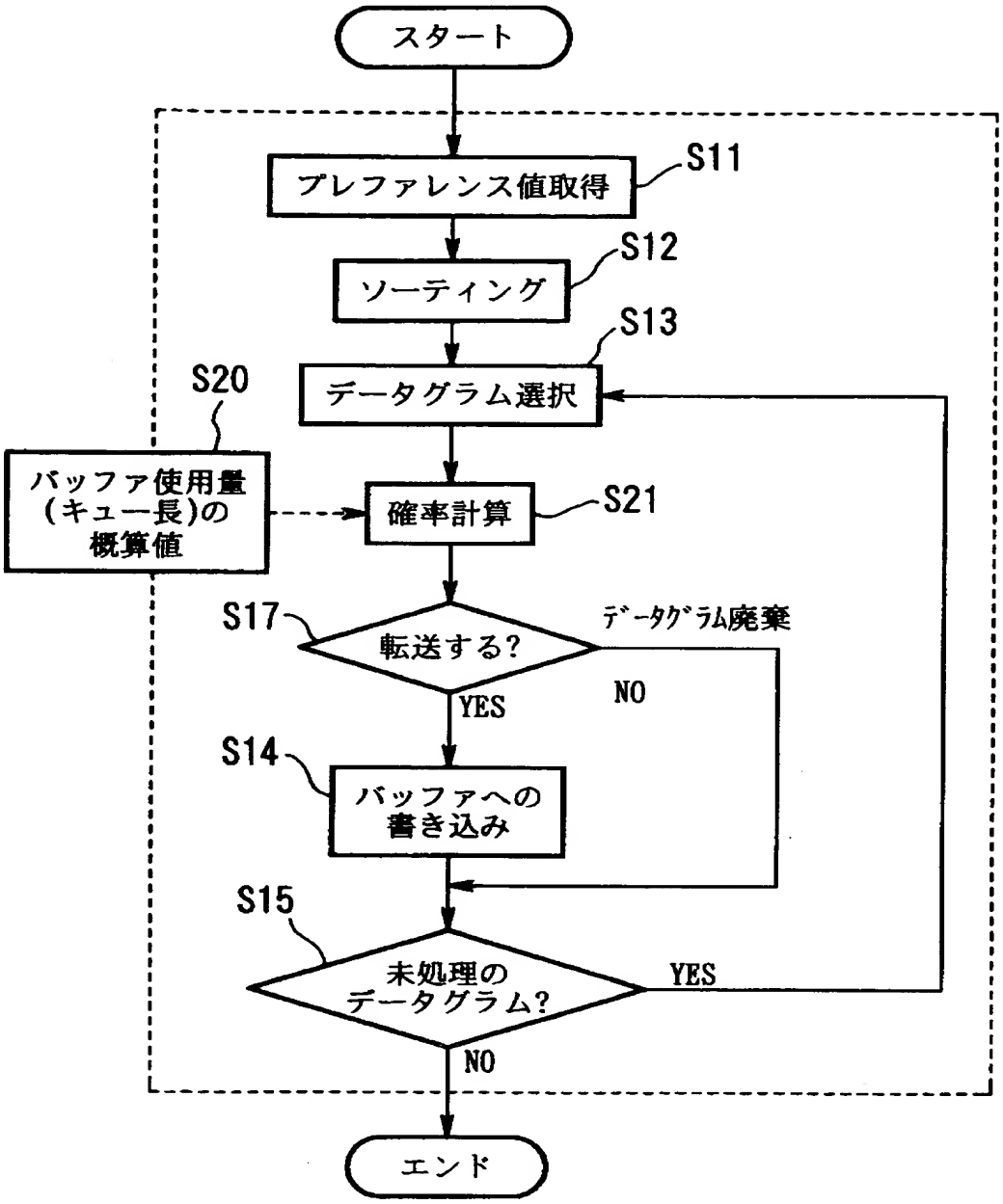
【図６】


スタート
S11
プレファレンス値取得
S12
ソーティング
S13
データグラム選択
S16
転送判断処理
S17
転送する?
データグラム廃棄
NO
YES
S14
バッファへの
書き込み
S15
未処理の
データグラム?
YES
NO
エンド

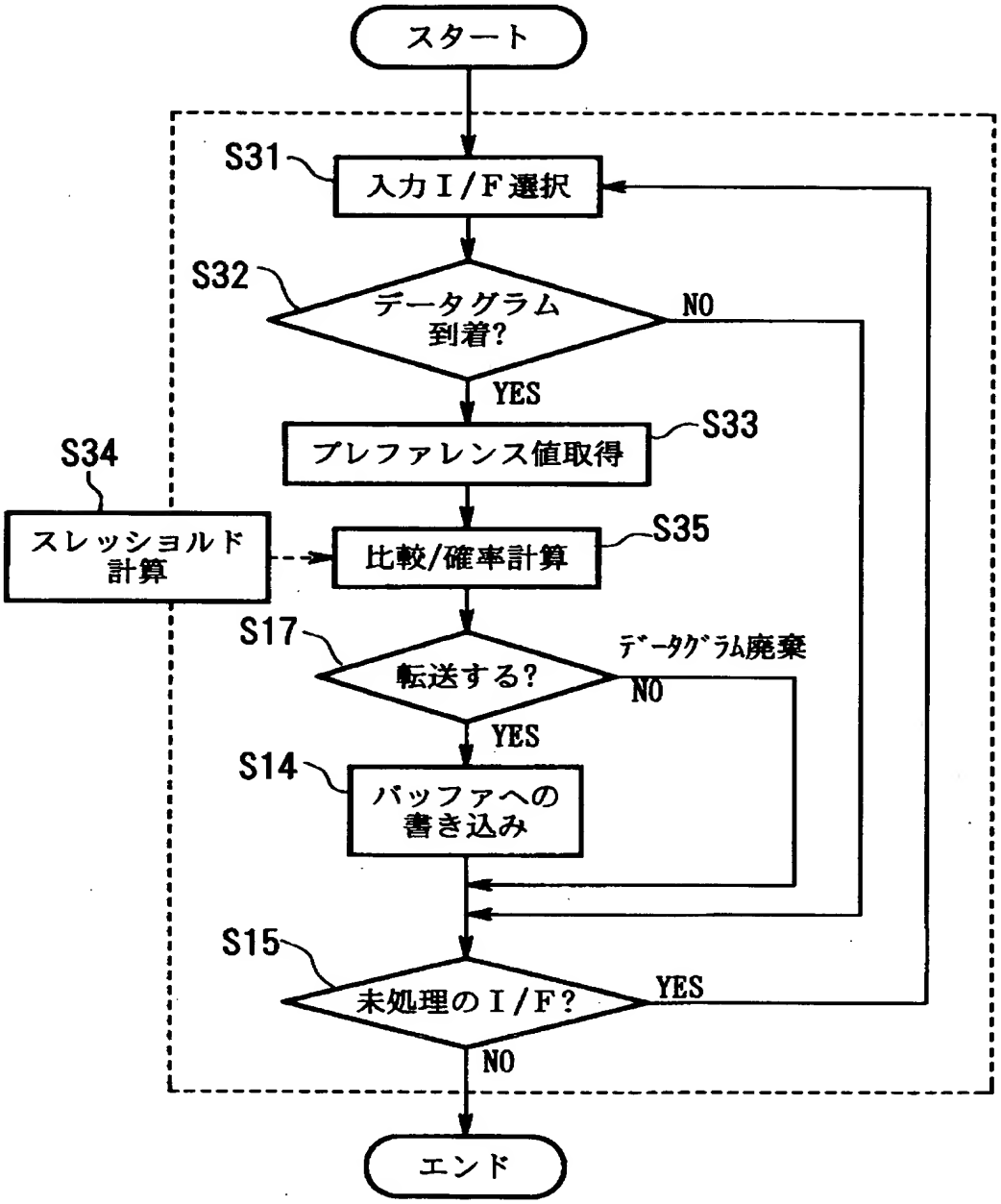
【図７】


スタート
プレファレンス値取得
S11
S12
ソーティング
S13
データグラム選択
S18
既処理済データグラムのプレファレンスの総和の計算
S19
確率計算
S17
転送する?
データグラム廃棄
NO
YES
S14
バッファへの書き込み
S15
未処理のデータグラム?
YES
NO
エンド

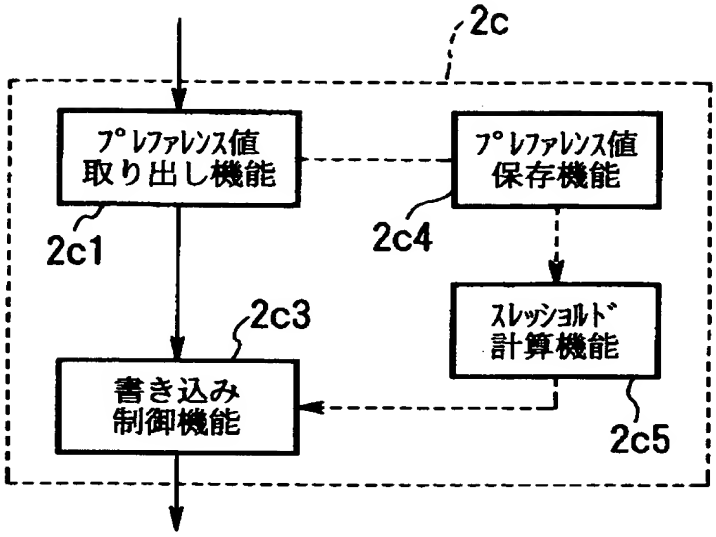
【図８】


スタート
プレファレンス値取得
S11
ソーティング
S12
S13
データグラム選択
S20
バッファ使用量
(キュー長)の
概算値
確率計算
S21
S17
転送する?
ﾃﾞｰﾀｸﾞﾗﾑ廃棄
NO
YES
S14
バッファへの
書き込み
S15
未処理の
データグラム?
YES
NO
エンド

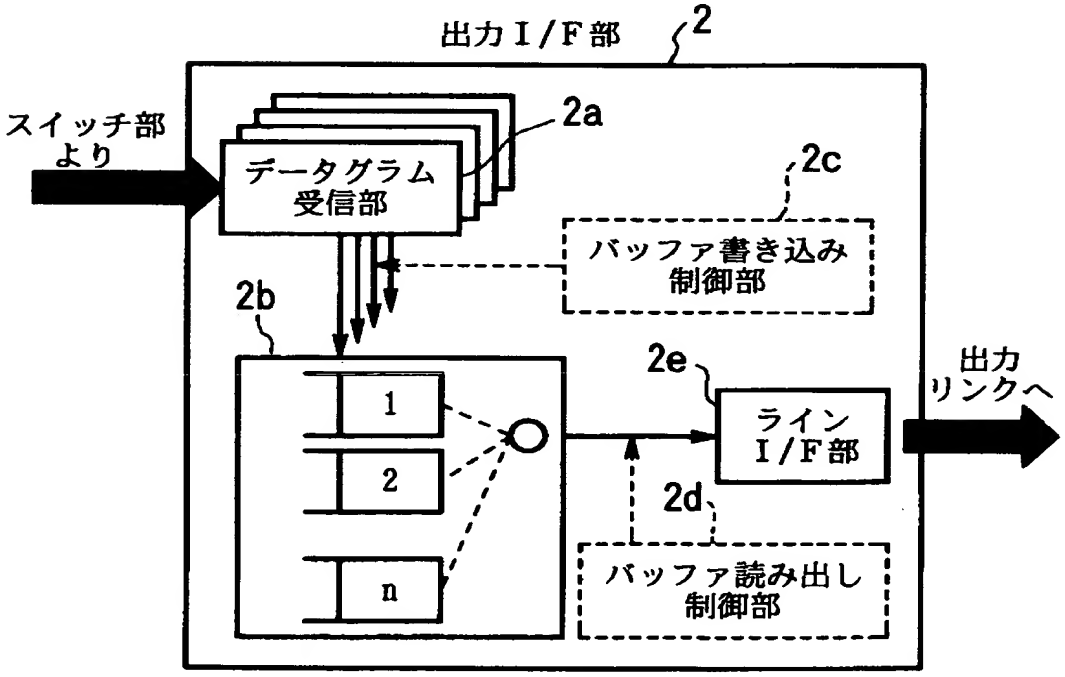
【図９】


スタート
S31
入力Ｉ/Ｆ選択
S32
データグラム
到着?
NO
YES
S33
プレファレンス値取得
S34
スレッショルド
計算
S35
比較/確率計算
S17
転送する?
ﾃﾞｰﾀｸﾞﾗﾑ廃棄
NO
YES
S14
バッファへの
書き込み
S15
未処理のＩ/Ｆ?
YES
NO
エンド

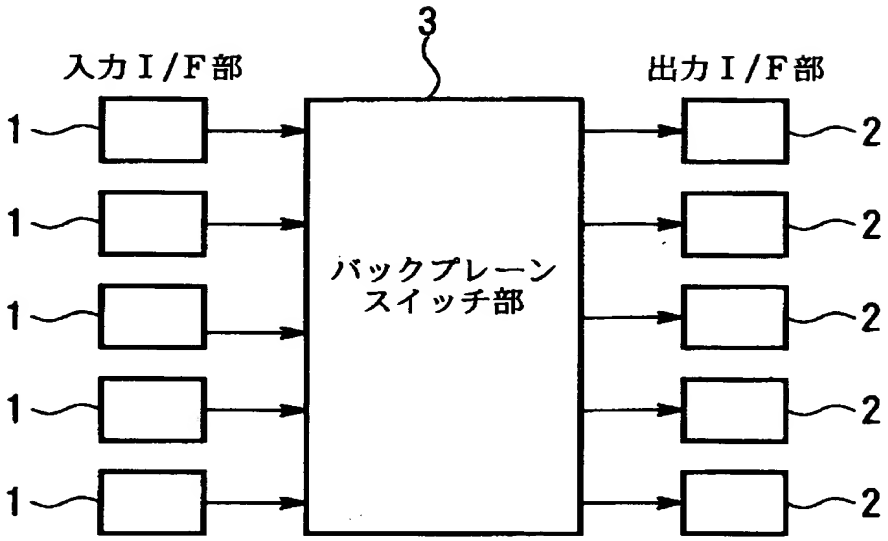
【図１０】


2c
ﾌﾟﾚﾌｧﾚﾝｽ値
取り出し機能
ﾌﾟﾚﾌｧﾚﾝｽ値
保存機能
2c1
2c4
2c3
ｽﾚｯｼｮﾙﾄﾞ
計算機能
書き込み
制御機能
2c5

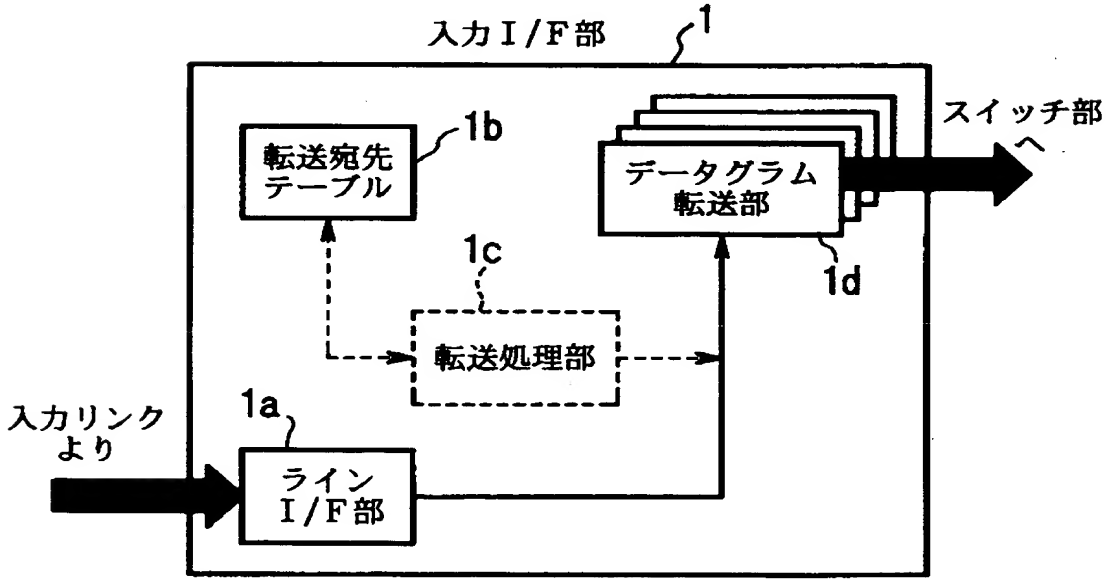
【図１１】


出力Ｉ/Ｆ部
2
スイッチ部
より
2a
データグラム
受信部
2c
バッファ書き込み
制御部
2b
2e
出力
リンクへ
1
2
n
ライン
Ｉ/Ｆ部
2d
バッファ読み出し
制御部

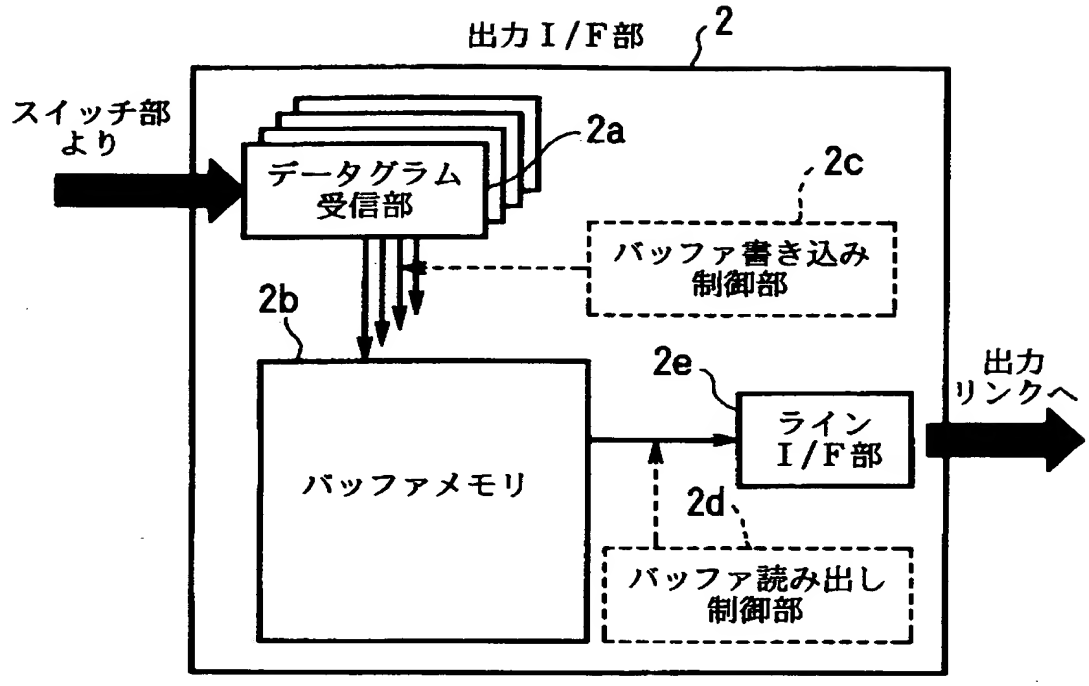
【図１２】


3
入力Ｉ/Ｆ部
出力Ｉ/Ｆ部
1
1
1
1
1
バックプレーン
スイッチ部
2
2
2
2
2

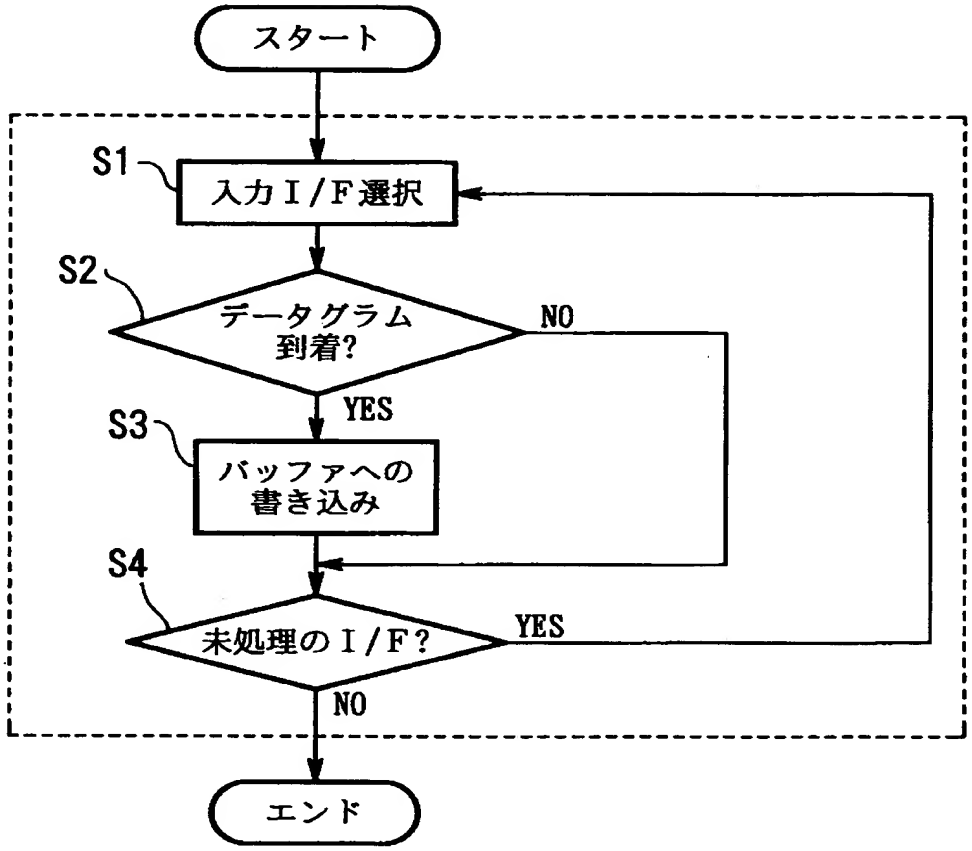
【図１３】


入力Ｉ/Ｆ部
1
1b
転送宛先
テーブル
データグラム
転送部
スイッチ部
へ
1c
1d
転送処理部
入力リンク
より
1a
ライン
Ｉ/Ｆ部

【図１４】

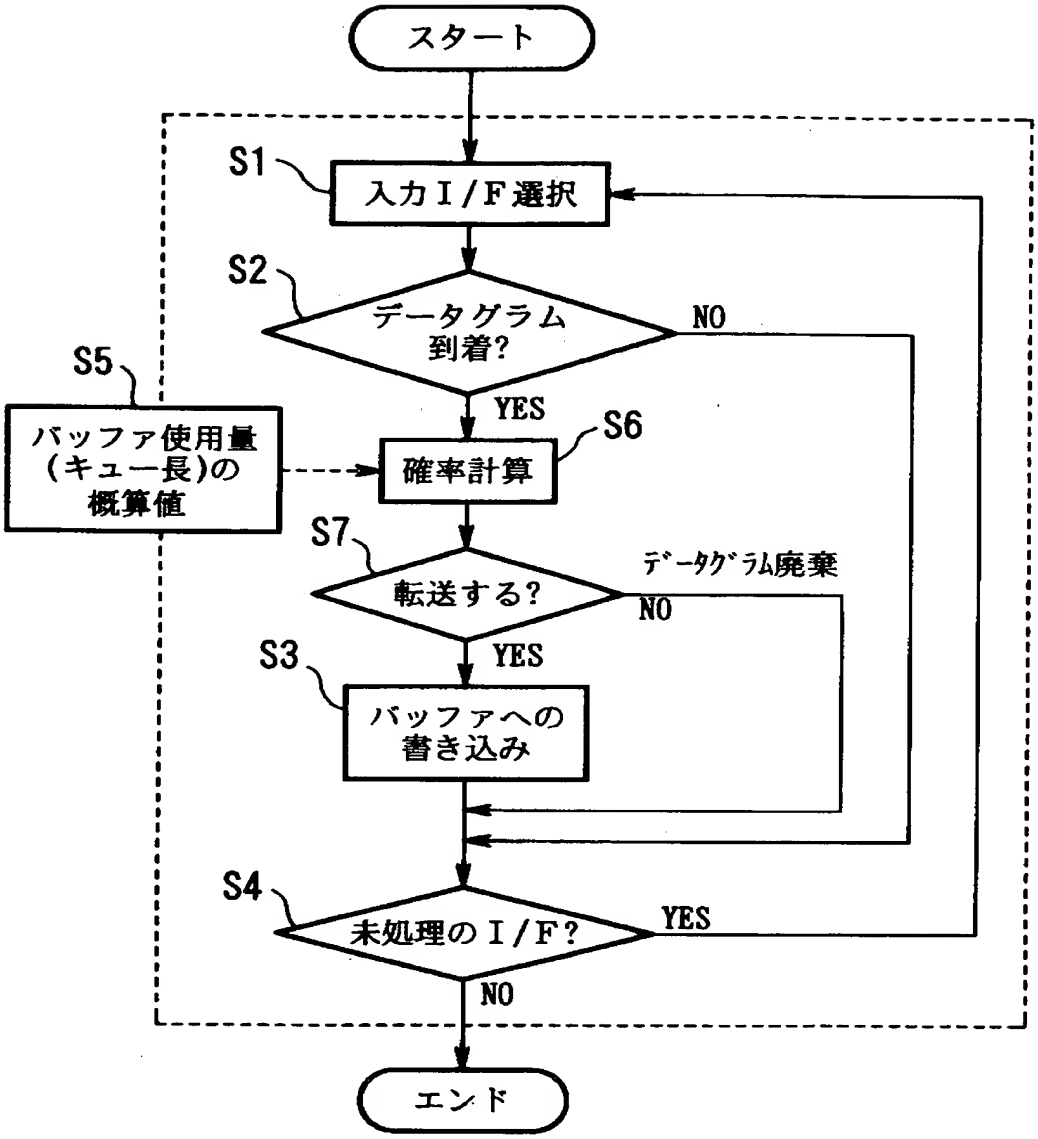

出力Ｉ/Ｆ部
2
スイッチ部
より
データグラム
受信部
2a
2c
バッファ書き込み
制御部
2b
2e
出力
リンクへ
ライン
Ｉ/Ｆ部
バッファメモリ
2d
バッファ読み出し
制御部

【図１５】



出証特２０００－３００２４５７

【図１６】


スタート
S1
入力Ｉ/Ｆ選択
S2
データグラム
到着?
NO
YES
S5
バッファ使用量
(キュー長)の
概算値
S6
確率計算
S7
転送する?
ﾃﾞｰﾀｸﾞﾗﾑ廃棄
NO
YES
S3
バッファへの
書き込み
S4
未処理のＩ/Ｆ?
YES
NO
エンド

【書類名】　　要約書

【要約】

【課題】　データグラムの送出に対する網の評価により得られる情報を利用して転送処理を行うことで、輻輳崩壊状態に陥らない安定したデータグラム通信網を実現可能にする。

【解決手段】　データグラムのヘッダに記載された転送宛先アドレスに向ってデータグラムをリレーするデータグラム転送システムにあって、データグラムに関するトラヒック情報に基づいてユーザのトラヒックの網に対するインパクトを評価し、その評価値をヘッダに書き込んで、この評価に従った順序による転送処理を行わせるトラヒック観測手段８を設ける。

【選択図】　　図１

出　願　人　履　歴　情　報

識別番号　　〔０００００４２２６〕

１．変更年月日　　１９９５年　９月２１日
［変更理由］　住所変更
住　所　東京都新宿区西新宿三丁目１９番２号
氏　名　日本電信電話株式会社

２．変更年月日　　１９９９年　７月１５日
［変更理由］　住所変更
住　所　東京都千代田区大手町二丁目３番１号
氏　名　日本電信電話株式会社